

Graph approach for observability analysis in power system dynamic state estimation

Akhila Kandivalasa, *Student Member, IEEE*, Marcos Netto, *Senior Member, IEEE*

Abstract—The proposed approach yields a numerical method that provably executes in linear time with respect to the number of nodes and edges in a graph. The graph, constructed from the power system model, requires only knowledge of the dependencies between state-to-state and output-to-state variables within a state-space framework. While *graph-based* observability analysis methods exist for power system static-state estimation, the approach presented here is the first for dynamic-state estimation (DSE). We examine decentralized and centralized DSE scenarios and compare our findings with a well-established, albeit non-scalable, observability analysis method in the literature. When compared to the latter in a centralized DSE setting, our method reduced computation time by 1440x.

Index Terms—Dynamic state estimation, graph theory, Kalman filter, Lie derivative, nonlinear dynamical systems, observability.

I. INTRODUCTION

Observability refers to the ability to determine a system's state from measurements. While well-understood for power system static state estimation [1], observability analysis is an active research area within dynamic state estimation (DSE) [2], with several challenges still to be addressed.

A grand challenge stems from the computational complexity of the numerical methods used for observability analysis in dynamic state estimation. The mainstream approach relies on Lie differentiation [3]. In what follows, for simplicity, we will refer to the class of methods based on Lie derivatives as the $\mathcal{L}$ approach. The $\mathcal{L}$ approach does not scale effectively [4], as noted by power system researchers [5]. Indeed, we confirmed this limitation. Our implementation of the Western System Coordinating Council (WSCC) test case (Example 7.1 in [6]) takes about 2 hours to run in MATLAB. Note that this test case has only twenty-one state variables. Although implementation details and computing power may differ, it is indisputable that the $\mathcal{L}$ approach is impractical for power system operators, as their systems often have thousands of state variables.

The empirical observability Gramian (EOG) [7] might serve as an alternative to the $\mathcal{L}$ approach. Unlike the $\mathcal{L}$ approach, which relies on *symbolic* computation of Lie derivatives, the EOG provides a way to evaluate observability *numerically*. One constructs the EOG by (i) applying small perturbations to the system's initial states, (ii) simulating the corresponding output trajectories over a finite horizon, and (iii) assembling these sensitivities into a symmetric matrix. This symmetric matrix corresponds to the Fisher information matrix, and its determinant (the product of its eigenvalues) measures the information the outputs carry about the initial state, and can thus be used to assess observability [8]. However, this reliance on perturbing the system and collecting data—which makes it *empirical*—introduces a key disadvantage. This approach

requires recomputation and additional simulations across many operating conditions, time windows, and outputs [7]. Repeated numerical simulations make the process computationally intensive, limiting its scalability to large-scale power grids. Hence, this approach does not solve the challenge of computational complexity.

Since the EOG and $\mathcal{L}$ approaches do not scale, an alternative could be to perform an observability analysis on a linear approximation of the original nonlinear system. However, relying on a linear approximation might lead to incorrect conclusions, as Rouhani and Abur [5] show using a *linearized* fourth-order synchronous machine model. They create an ideal scenario where all state variables are assumed to be measured and a second scenario where only a single reactive power measurement is available. One would expect the ideal scenario to yield stronger observability, since all state variables are directly measured. Yet, because of linearization, observability analysis favors the second (non-ideal) scenario [5].

This paper departs from the EOG and $\mathcal{L}$ approaches, and does not rely on linearization. We introduce a method for analyzing observability that uses a directed graph (digraph) constructed from the nonlinear differential equations [9] that model power systems. From now on, we refer to the proposed method as the *graph approach*. We show that the graph approach produces results equivalent to those of the $\mathcal{L}$ approach but needs significantly less computational effort. Specifically, for the WSCC test case with the same conditions used to evaluate the $\mathcal{L}$ approach, it takes less than 5 seconds to check for observability using the proposed method. At the core of this method is Kosaraju-Sharir's algorithm, which finds all *strongly-connected components* (SCCs)—defined in Section IV and central to the proposed method—in a digraph and operates in linear time relative to the number of nodes and edges [10]. Therefore, the proposed method can scale to bulk power grids. While there are numerical and graph-based (topological) observability analysis methods for power system static state estimation [11], this is the first time a graph-based method has been proposed for observability analysis in power system *dynamic state estimation* [2].

This paper proceeds as follows. Section II briefly presents the $\mathcal{L}$ approach to observability analysis. Section III introduces the proposed graph approach. Section IV demonstrates the graph approach for both decentralized and centralized dynamic state estimation. Section V summarizes the key findings and outlines future work.

II. THE $\mathcal{L}$ APPROACH TO OBSERVABILITY ANALYSIS

Consider a nonlinear dynamical system,

$$\dot{x}(t) = f(x(t), u(t)), \quad y(t) = h(x(t)) \quad (1)$$

evolving on a smooth manifold $\mathcal{M} \subseteq \mathbb{R}^n$, where the state $x \in \mathcal{M}$, the input $u \in \mathbb{R}^m$, and the output $y \in \mathbb{R}^p$; f defines the system dynamics, and h relates measured outputs with

This work is supported by the National Science Foundation under Grant 2328241. The authors are with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA.

state variables. The notion of indistinguishability [3] provides the basis for defining observability in dynamical systems.

Definition 1. Let $\mathbf{x}(0) =: \mathbf{x}_0$ denote a possible state of (1) at $t = 0$. Two states $\mathbf{x}_0, \mathbf{x}'_0 \in \mathcal{M}$ are *indistinguishable* if for every admissible input $\mathbf{u}(t)$, the resulting output trajectories are identical over a finite time interval, i.e., $\mathbf{h}(\mathbf{x}(t; \mathbf{x}_0, \mathbf{u})) = \mathbf{h}(\mathbf{x}(t; \mathbf{x}'_0, \mathbf{u}))$ for all $\mathbf{u}(t)$, $t \in [t_0, t_1]$.

The set of all such indistinguishable states from $\mathbf{x}_0$ is denoted as $\mathcal{I}(\mathbf{x}_0) = \{\mathbf{x}_0, \mathbf{x}'_0\}$.

Definition 2. A system (1) is *observable at $\mathbf{x}_0$* if the only state indistinguishable from $\mathbf{x}_0$ is $\mathbf{x}_0$ itself, that is, $\mathcal{I}(\mathbf{x}_0) = \{\mathbf{x}_0\}$.

While precise, Definition 2 is global in nature and requires knowledge of the entire space of input–output trajectories over the considered time interval. Such a condition is challenging to verify numerically for nonlinear systems. So, Hermann and Krener [3] proceed by defining indistinguishability locally over a small neighborhood $\mathcal{V}$ in an open neighborhood $\mathcal{U}$ of $\mathbf{x}$, $\mathcal{V} \subset \mathcal{U}$, from which another definition for observability emerges.

Definition 3. A system (1) is *locally weakly observable at $\mathbf{x}_0$* if $\mathcal{I}_{\mathcal{V}}(\mathbf{x}_0) = \{\mathbf{x}_0\}$, and is *locally weakly observable* if it is so at every $\mathbf{x} \in \mathcal{M}$.

Simply put, based on Definition 3, a system is observable if no other state produces the same input–output trajectory within a small neighborhood of $\mathbf{x}$. Definition 3 is amenable to a numerical test to check for observability—the $\mathcal{L}$ approach, which we introduce next. We refer the interested reader to [3] for a more formal treatment and further details on observability definitions. Now, consider a map

$$\mathcal{G} : \mathbf{x} \mapsto [\mathbf{y} \quad \mathbf{y}^{(1)} \quad \mathbf{y}^{(2)} \quad \dots \quad \mathbf{y}^{(d)}]^\top \quad (2)$$

where $\mathbf{y}^{(k)}$ denotes the k th-order derivative of $\mathbf{y}$, $k \in \mathbb{N}$. If $\mathcal{G}$ is invertible for some $d \leq (n-1)$, then one can reconstruct the state from measurements; hence, the system is observable. If we consider Definition 3, then the Implicit Function Theorem provides a sufficient condition for local invertibility. The map $\mathcal{G}$ is locally invertible at $\mathbf{x}_0$ if the Jacobian

$$\mathcal{O} := \left. \frac{\partial \mathcal{G}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_0} \quad (3)$$

referred to as the *observability matrix* has full rank, i.e., if

$$\text{rank} \left(\left. \frac{\partial \mathcal{G}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_0} \right) = n. \quad (4)$$

If (4) holds, (1) is locally weakly observable. Rouhani and Abur [5] apply condition (4). In what follows, for simplicity, a system satisfying condition (4) is said *observable*, although the $\mathcal{L}$ approach derives from Definition 3. Now, the Lie derivatives of $\mathbf{y}$ along the vector field $\mathbf{f}$

$$\mathcal{L}_{\mathbf{f}} \mathbf{y} := \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_1} \dot{x}_1 + \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_2} \dot{x}_2 + \dots + \frac{\partial \mathbf{h}(\mathbf{x})}{\partial x_n} \dot{x}_n \quad (5)$$

measure how outputs $\{y_1, \dots, y_p\}$ in $\mathbf{y}$ change along the vector field $\mathbf{f}$. Higher-order Lie derivatives are given by

$$\mathcal{L}_{\mathbf{f}}^k \mathbf{y} = \frac{\partial \mathcal{L}_{\mathbf{f}}^{k-1} \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}, \quad 1 \leq k \leq n, \quad (6)$$

and the zeroth order $\mathcal{L}_{\mathbf{f}}^0 \mathbf{y} := \mathbf{h}(\mathbf{x})$. With these definitions, the map (2) becomes

$$\mathbf{g} = [\mathcal{L}_{\mathbf{f}}^0 y_1 \quad \dots \quad \mathcal{L}_{\mathbf{f}}^{n-1} y_1 \quad \mathcal{L}_{\mathbf{f}}^0 y_2 \quad \dots \quad \mathcal{L}_{\mathbf{f}}^{n-1} y_p]^\top \quad (7)$$

and the observability matrix $\mathcal{O} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_0}$.

Example 1. Let a nonlinear dynamical system be

$$\dot{x}_1 = \sin(x_2) - x_1 + x_4, \quad \dot{x}_2 = x_3, \quad \dot{x}_3 = x_1^2, \quad \dot{x}_4 = x_4 \quad (8)$$

and consider three measurement models: (i) $y = h(\mathbf{x}) = x_2$ (ii) $y' = h(\mathbf{x}) = x_2 + \sin(x_1)$ and (iii) $y'' = h(\mathbf{x}) = x_4$. For case (i), $\mathcal{L}_{\mathbf{f}}^0 y = x_2$, and

$$\begin{aligned} \mathcal{L}_{\mathbf{f}}^1 y &= \frac{\partial \mathcal{L}_{\mathbf{f}}^0 y}{\partial \mathbf{x}} \mathbf{f} = \frac{\partial x_2}{\partial x_1} \dot{x}_1 + \frac{\partial x_2}{\partial x_2} \dot{x}_2 + \frac{\partial x_2}{\partial x_3} \dot{x}_3 + \frac{\partial x_2}{\partial x_4} \dot{x}_4 \\ &= 0 \cdot \dot{x}_1 + 1 \cdot \dot{x}_2 + 0 \cdot \dot{x}_3 + 0 \cdot \dot{x}_4 = \dot{x}_2 = x_3 \\ \mathcal{L}_{\mathbf{f}}^2 y &= \frac{\partial \mathcal{L}_{\mathbf{f}}^1 y}{\partial \mathbf{x}} \mathbf{f} = \frac{\partial x_3}{\partial x_1} \dot{x}_1 + \frac{\partial x_3}{\partial x_2} \dot{x}_2 + \frac{\partial x_3}{\partial x_3} \dot{x}_3 + \frac{\partial x_3}{\partial x_4} \dot{x}_4 \\ &= 0 \cdot \dot{x}_1 + 0 \cdot \dot{x}_2 + 1 \cdot \dot{x}_3 + 0 \cdot \dot{x}_4 = \dot{x}_3 = x_1^2 \\ \mathcal{L}_{\mathbf{f}}^3 y &= \frac{\partial \mathcal{L}_{\mathbf{f}}^2 y}{\partial \mathbf{x}} \mathbf{f} = \frac{\partial x_1^2}{\partial x_1} \dot{x}_1 + \frac{\partial x_1^2}{\partial x_2} \dot{x}_2 + \frac{\partial x_1^2}{\partial x_3} \dot{x}_3 + \frac{\partial x_1^2}{\partial x_4} \dot{x}_4 \\ &= 2x_1 \dot{x}_1 = 2x_1 (\sin x_2 - x_1 + x_4) \end{aligned}$$

Note that applying higher-order Lie derivatives is “like peeling off layers to reveal hidden dependencies between outputs and state variables, including those not explicit in the original output function.” The observability matrix for case (i) is

$$\mathbf{O}_1(\mathbf{x}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2x_1 & 0 & 0 & 0 \\ 2 \sin x_2 - 4x_1 + 2x_4 & 2x_1 \cos x_2 & 0 & 2x_1 \end{bmatrix}.$$

Under generic conditions such as $x_1 \neq 0$, the rows of $\mathbf{O}_1(\mathbf{x})$ are linearly independent, and the matrix attains full rank. The system is, therefore, observable. The observability matrix for case (ii) is

$$\mathbf{O}_2(\mathbf{x}) = \begin{bmatrix} \cos x_1 & 1 & 0 & 0 \\ \star_{2,1}(\mathbf{x}) & \star_{2,2}(\mathbf{x}) & \star_{2,3}(\mathbf{x}) & \star_{2,4}(\mathbf{x}) \\ \star_{3,1}(\mathbf{x}) & \star_{3,2}(\mathbf{x}) & \star_{3,3}(\mathbf{x}) & \star_{3,4}(\mathbf{x}) \\ \star_{4,1}(\mathbf{x}) & \star_{4,2}(\mathbf{x}) & \star_{4,3}(\mathbf{x}) & \star_{4,4}(\mathbf{x}) \end{bmatrix}.$$

Case (ii) yields long symbolic expressions. For instance,

$$\begin{aligned} \star_{3,2}(\mathbf{x}) &= -\cos(x_2)(\cos(x_1) + \sin(x_1)(x_4 - x_1 + \sin(x_2))) \\ &\quad - x_3 \cos(x_1) \sin(x_2) - \cos(x_2) \sin(x_1)(x_4 - x_1 + \sin(x_2)) \end{aligned}$$

For simplicity, we denote those expressions by $\star_{\text{row}, \text{col}}(\mathbf{x})$ in $\mathbf{O}_2$. It can be verified that $\mathbf{O}_2$ is generically of full rank, hence the system is observable. For case (iii), $\mathcal{L}_{\mathbf{f}}^0 y' = x_4$, and

$$\begin{aligned} \mathcal{L}_{\mathbf{f}}^1 y' &= 5x_4 \\ \mathcal{L}_{\mathbf{f}}^2 y' &= 25x_4 \\ \mathcal{L}_{\mathbf{f}}^3 y' &= 125x_4 \end{aligned} \quad \mathbf{O}_3(\mathbf{x}) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 25 \\ 0 & 0 & 0 & 125 \end{bmatrix}.$$

Clearly, $\mathbf{O}_3$ is rank deficient, and the system is not observable. $\square$

In general, $\dim(\mathcal{O}) = np \times n$. However, the rank condition (4) may be satisfied before computing all Lie derivatives up to order $n-1$. In such cases, the hidden dependencies between outputs and state variables are unveiled at $k < n$. Therefore, the dimension of $\mathcal{O}$ is more accurately given by $(k+1)p \times n$, where $(k+1) \leq n$ [7]. Even when the rank condition is satisfied at a lower order $k < n$, the intermediate symbolic expressions associated with the Lie derivatives rapidly expand in size. This is true even for the simple system (8), as evidenced from the structure of $\mathbf{O}_2$. This makes the $\mathcal{L}$ approach computationally demanding, and applying it to multi-machine power systems is infeasible. Rouhani and Abur [5] apply the $\mathcal{L}$ approach to a fourth-order synchronous generator model with an IEEE type-1 exciter, a system with seven state variables. Section IV of this paper demonstrates the proposed graph approach using the same decentralized model in [5]. It further presents results for centralized dynamic state estimation.

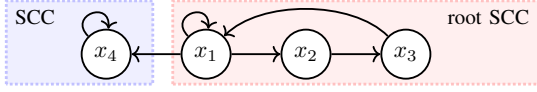


Fig. 1. Digraph of (8).

III. THE GRAPH APPROACH TO OBSERVABILITY ANALYSIS

Let a digraph built from (1) be given by $\mathcal{D} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of nodes, and $\mathcal{E}$ is a binary relation on $\mathcal{V}$. The elements of $\mathcal{E}$ are called edges. Each node corresponds to a state variable $x_i \in \mathcal{V}$, and a directed edge $x_i \rightarrow x_j$ exists whenever the differential equation $\dot{x}_i$ explicitly depends on x_j , i.e., $\partial \dot{x}_i / \partial x_j \neq 0$. As defined, $\mathcal{D}$ may include self-loops, i.e., $x_i \rightarrow x_i$ edges. The set of all such dependencies can be compactly represented in an adjacency matrix with entries $a_{ij} = 1$ if $(x_i \rightarrow x_j) \in \mathcal{E}$ and 0 otherwise.

Definition 4. A path $x_j - x_i$ is a sequence $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_k)$ with $\varepsilon_0 = x_j$ and $\varepsilon_k = x_i$, and $(\varepsilon_{\ell-1}, \varepsilon_\ell) \in \mathcal{E} \forall \ell \in \{\ell\}_{\ell=1}^k$.

Definition 5. A node x_i is *reachable* from x_j if there is a path from x_j to x_i .

Definition 6. A digraph $\mathcal{D}$ is *strongly connected* if every two nodes are reachable from each other.

Definition 7. A subgraph $\mathcal{D}_{\text{sub}} = (\mathcal{V}_{\text{sub}}, \mathcal{E}_{\text{sub}})$ of $\mathcal{D}$ satisfies $\mathcal{V}_{\text{sub}} \subseteq \mathcal{V}$ and $\mathcal{E}_{\text{sub}} \subseteq \mathcal{E}$, with $\mathcal{E}_{\text{sub}}$ containing ordered pairs.

Definition 8. *Strongly connected components* (SCCs) are the largest subgraphs, ordered by number of nodes, such that each component is a strongly connected graph.

Definition 9. A *root strongly connected component* (root SCC) is an SCC with no incoming edges.

Example 2. The digraph of (8) (Fig. 1) has six directed edges: $x_1 \rightarrow x_1$, $x_1 \rightarrow x_2$, $x_1 \rightarrow x_4$, $x_2 \rightarrow x_3$, $x_3 \rightarrow x_1$, and $x_4 \rightarrow x_4$. Nodes x_1 , x_2 , and x_3 form an SCC, since one can reach every node from every other node. Furthermore, since there are no incoming edges to this SCC, it is also a root SCC. $\square$

Root SCCs act as information sources within $\mathcal{D}$ because all nodes in $\mathcal{D}$ can be traced back to at least one node within a root SCC; however, the reverse is not true.

Proposition 1. A dynamical system (1) is termed *structurally observable* if, for a digraph $\mathcal{D}$ built from (1), at least one node from every root SCC is measured [9].

Proposition 1 can be interpreted as the structural counterpart of the algebraic rank condition used in the $\mathcal{L}$ approach, where the Lie-derivative expansion of the outputs captures the same dependencies in an analytical form. Proposition 1 offers a necessary and almost everywhere sufficient condition for observability [12]. Note that if the outputs y in (1) depend on at least one state variable present in every root SCC of $\mathcal{D}$, the system is observable.

Let $\mathcal{X}_r \subseteq \{x_1, \dots, x_n\}$ denote the subset of state variables in root SCCs of $\mathcal{D}$, and $\mathcal{X}_d = \{x_1, \dots, x_n\} - \mathcal{X}_r$ the set with possibly remaining state variables. Consider an output of the form $y = g(\mathcal{X}_r)$. A first-order Taylor series expansion of $g(\mathcal{X}_r)$ at $x = x_0$ is given by,

$$g(\mathcal{X}_r) = g(x_0) = c^\top x_0 \quad (9)$$

where c denotes a vector of constants. Thus, y provides direct access to the state variables in $\mathcal{X}_r$. Note that this is consistent with the $\mathcal{L}$ approach being defined locally, cf. Definition 3. Measuring any function of root-SCC variables injects information from all independent sources; this information is carried through the directed edges and reaches any remaining state variables in $\mathcal{X}_d$. As the graph approach is based on dependencies in the system equations, it naturally extends

Algorithm 1: Extraction of SCCs and Root SCCs from Adjacency Matrix

Require: Adjacency matrix $A = [a_{ij}] \in \{0, 1\}^{n \times n}$
Ensure : SCCs, $\mathcal{C} = \{C_1, \dots, C_m\}$, and root SCCs, $\mathcal{R}$

- 1 **Step 1: Compute finishing order of nodes;**
- 2 $\text{visit} \leftarrow [0, \dots, 0]_{1 \times n}$, $\text{stack} \leftarrow \text{empty}$;
- 3 **for** $i \leftarrow 1$ **to** n **do**
- 4 **if** $\text{visit}[i] = 0$ **then**
- 5 $[\text{visit}, \text{stack}] \leftarrow \text{FillOrder}(i, A, \text{visit}, \text{stack})$;
- 6 **end**
- 7 **end**
- 8 **Procedure** $\text{FillOrder}(u, A, \text{visit}, \text{stack})$:
- 9 $\text{visit}[u] \leftarrow 1$;
- 10 **for** $v \leftarrow 1$ **to** n **do**
- 11 **if** $A_{uv} = 1$ **and** $\text{visit}[v] = 0$ **then**
- 12 $[\text{visit}, \text{stack}] \leftarrow \text{FillOrder}(v, A, \text{visit}, \text{stack})$;
- 13 **end**
- 14 **end**
- 15 $\text{append } u \text{ to stack}$;
- 16 **Step 2: Identify SCCs on transposed graph;**
- 17 $A^\top = \text{transpose}(A)$; $\text{visit} \leftarrow [0, \dots, 0]_{1 \times n}$; $\mathcal{C} \leftarrow \text{empty}$;
- 18 **while** stack not empty **do**
- 19 $u \leftarrow \text{stack}[\text{end}]$; $\text{stack}[\text{end}] = []$;
- 20 **if** $\text{visit}[u] = 0$ **then**
- 21 $\text{comp} \leftarrow \text{empty}$;
- 22 $[\text{visit}, \text{comp}] \leftarrow \text{dfs}(u, A^\top, \text{visit}, \text{comp})$;
- 23 $\text{append sorted}(\text{comp}) \text{ to } \mathcal{C}$;
- 24 **end**
- 25 **end**
- 26 **Procedure** $\text{dfs}(u, A^\top, \text{visit}, \text{comp})$:
- 27 $\text{visit}[u] \leftarrow 1$; $\text{append } u \text{ to comp}$;
- 28 **for** $v \leftarrow 1$ **to** n **do**
- 29 **if** $A_{uv}^\top = 1$ **and** $\text{visit}[v] = 0$ **then**
- 30 $[\text{visit}, \text{comp}] \leftarrow \text{dfs}(v, A^\top, \text{visit}, \text{comp})$;
- 31 **end**
- 32 **end**
- 33 **Step 3: Identify Root SCCs;**
- 34 $m \leftarrow \text{length}(\mathcal{C})$; $\text{root_flags} \leftarrow [1, 1, \dots, 1]_{1 \times m}$;
- 35 **for** $i \leftarrow 1$ **to** m **do**
- 36 **for** $j \leftarrow 1$ **to** m **do**
- 37 **if** $i \neq j$ **then**
- 38 **for each** $u \in C_j$ **do**
- 39 **for each** $v \in C_i$ **do**
- 40 **if** $A_{uv} = 1$ **then**
- 41 $\text{root_flags}[i] \leftarrow 0$;
- 42 **end**
- 43 **end**
- 44 **end**
- 45 **end**
- 46 **end**
- 47 **end**
- 48 $\mathcal{R} \leftarrow \{C_i \text{ whose } \text{root_flags}[i] = 1\}$;
- 49 **Output:** List of SCCs and list of root SCCs.

to DAE formulations, where algebraic equations also induce dependencies. Thus, for DAE formulations, more nodes are introduced in the digraph, and correspondingly, more rows and columns are added to the adjacency matrix. Algorithm 1 extracts SCCs and root SCCs from an adjacency matrix built from (1). The procedure relies on Kosaraju-Sharir's algorithm,

TABLE I
DEFINITION OF CONSTANTS IN (10)

$c_{1i} = 1/T'_{d0i}$	$c_{6i} = \omega_s/(2H_i)$	$c_{11i} = 1/T_{Fi}$
$c_{2i} = X'_{di} - X'_{di}$	$c_{7i} = X'_{qi} - X'_{di}$	$c_{12i} = K_{Fi}/T_{Fi}$
$c_{3i} = 1/T'_{q0i}$	$c_{8i} = D_i$	$c_{13i} = 1/T_{Ai}$
$c_{4i} = X'_{qi} - X'_{qi}$	$c_{9i} = 1/T_{Ei}$	$c_{14i} = K_{Ai}$
$c_{5i} = \omega_s$	$c_{10i} = K_{Ei}$	$c_{15i} = K_{Ai}K_{Fi}/T_{Fi}$

whose computational complexity $\mathcal{O}(|\mathcal{V}| + |\mathcal{E}|)$ scales linearly with the number of vertices and edges in a digraph [10].

We are now able to present results on realistic models of power systems.

IV. NUMERICAL RESULTS

We consider the Example 7.1 in [6], a three-machine ($m=3$) nine-bus ($b=9$) system where each machine model,

$$\dot{E}'_{qi} = c_{1i} (-E'_{qi} - c_{2i}I_{di} + E_{fdi}) \quad i = 1, \dots, m \quad (10)$$

$$\dot{E}'_{di} = c_{3i} (-E'_{di} + c_{4i}I_{qi})$$

$$\dot{\delta}_i = \omega_i - c_{5i}$$

$$\dot{\omega}_i = c_{6i} (T_{Mi} - E'_{di}I_{di} - E'_{qi}I_{qi} - c_{7i}I_{di}I_{qi} - c_{8i}(\omega_i - c_{5i}))$$

$$\dot{E}_{fdi} = c_{9i} (- (c_{10i} + S_{Ei}(E_{fdi})) E_{fdi} + V_{Ri})$$

$$\dot{R}_{fi} = c_{11i} (-R_{fi} + c_{12i}E_{fdi})$$

$$\dot{V}_{Ri} = c_{13i} (-V_{Ri} + c_{14i}R_{fi} - c_{15i}E_{fdi} + c_{14i}(V_{refi} - V_i))$$

connects to the network via algebraic constraints, as follows. For each machine bus $i = 1, \dots, m$, there are two stator voltage balance equations,

$$E'_{di} - V_i \sin(\delta_i - \theta_i) - R_{si}I_{di} + X'_{qi}I_{qi} = 0 \quad (11)$$

$$E'_{qi} - V_i \cos(\delta_i - \theta_i) - R_{si}I_{qi} - X'_{di}I_{di} = 0 \quad (12)$$

$V_i e^{j\theta_i} = V_i \cos \theta_i + jV_i \sin \theta_i = V_{Di} + jV_{Qi}$, and two network equations (presented below in polar form as a single equation)

$$0 = V_i e^{j\theta_i} (I_{di} - jI_{qi}) e^{-j(\delta_i - \pi/2)} - \sum_{k=1}^n V_i V_k Y_{ik} e^{j(\theta_i - \theta_k - \alpha_{ik})} + P_{Li} + jQ_{Li}. \quad (13)$$

For each load bus $i = m+1, \dots, b$, two network equations

$$0 = - \sum_{k=1}^n V_i V_k Y_{ik} e^{j(\theta_i - \theta_k - \alpha_{ik})} + P_{Li} + jQ_{Li}. \quad (14)$$

This model (10)–(14) is well-known; see [6] for further details. Note that the graph approach is parameter-agnostic.

Case 1: Decentralized dynamic state estimation

Following [13], we treat voltages as inputs and currents as outputs to separate the dynamics of each generator from the rest of the network without ignoring their interaction. Thus,

$$\mathbf{x} = [E'_{qi} \ E'_{di} \ \delta_i \ \omega_i \ E_{fdi} \ R_{fi} \ V_{Ri}]^T, \quad (15)$$

$$\mathbf{u} = [T_{Mi} \ V_{refi} \ V_{Di} \ V_{Qi}]^T, \text{ and } \mathbf{y} = [I_{Di} \ I_{Qi}]^T,$$

where

$$I_{Di} + jI_{Qi} = (I_{di} + jI_{qi}) e^{j(\delta_i - \frac{\pi}{2})} \quad (16)$$

This particular choice of $\{V_{Di}, V_{Qi}, I_{Di}, I_{Qi}\}$ is consistent with the availability of synchrophasor measurements at the terminal of each machine. Now, following [6], we define

$$\mathbf{Z}_{dqi} \triangleq \begin{bmatrix} R_{si} & -X'_{qi} \\ X'_{di} & R_{si} \end{bmatrix}. \quad (17)$$

Using (17) to algebraically manipulate (11)–(12) yields

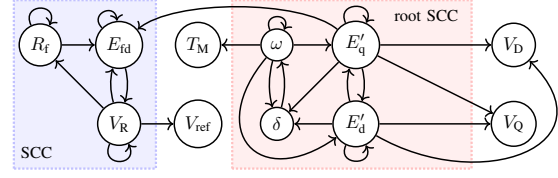


Fig. 2. Digraph for Case 1: Decentralized dynamic state estimation.

$$\begin{bmatrix} I_{di} \\ I_{qi} \end{bmatrix} = [\mathbf{Z}_{dqi}]^{-1} \begin{bmatrix} E'_{di} - V_i \sin(\delta_i - \theta_i) \\ E'_{qi} - V_i \cos(\delta_i - \theta_i) \end{bmatrix}, \quad i = 1, \dots, m. \quad (18)$$

From (16) and (18), it is evident that I_{Di} and I_{Qi} are functions of the state variables $\{E'_{qi}, E'_{di}, \delta_i\}$. Next, because the graph approach requires only knowledge of the dependencies in a state-space model, we will use the simplified notation

$$\begin{aligned} I_{Di} &= h_1(\{\mathcal{X}\}, \{\mathcal{U}\}) = h_1(\{E'_{qi}, E'_{di}, \delta_i\}, \{V_{Di}, V_{Qi}\}) \\ I_{Qi} &= h_2(\{E'_{qi}, E'_{di}, \delta_i\}, \{V_{Di}, V_{Qi}\}) \end{aligned} \quad (19)$$

where $\mathcal{X}$ (respectively, $\mathcal{U}$) denotes a subset of the variables in $\mathbf{x}$ (respectively, $\mathbf{u}$), cf. (1). Substituting I_{Di} and I_{Qi} in (10),

$$\begin{aligned} \dot{E}'_{qi} &= f_1(\{E'_{qi}, E'_{di}, \delta_i, E_{fdi}\}, \{V_{Di}, V_{Qi}\}) \\ \dot{E}'_{di} &= f_2(\{E'_{qi}, E'_{di}, \delta_i\}, \{V_{Di}, V_{Qi}\}) \\ \dot{\delta}_i &= f_3(\{\omega_i\}, \{\}) \\ \dot{\omega}_i &= f_4(\{E'_{qi}, E'_{di}, \delta_i, \omega_i\}, \{T_{Mi}\}) \\ \dot{E}_{fdi} &= f_5(\{E_{fdi}, V_{Ri}\}, \{\}) \\ \dot{R}_{fi} &= f_6(\{E_{fdi}, R_{fi}\}, \{\}) \\ \dot{V}_{Ri} &= f_7(\{E_{fdi}, R_{fi}, V_{Ri}\}, \{V_{refi}\}). \end{aligned} \quad (20)$$

From (20), the adjacency matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad (21)$$

where each row is associated with a differential equation $\dot{x}_i$, in the order they appear in (20), and each column with a state variable x_i . An entry a_{ij} in (21) is assigned the value ‘1’ if the differential equation includes x_j (i.e., $\frac{\partial \dot{x}_i}{\partial x_j} \neq 0$), and zero otherwise; zeros are omitted in (21) for clarity. We apply a modified Kosaraju-Sharir’s Algorithm 1 [10] to the adjacency matrix to identify strongly connected components (SCCs) and systematically determine the root SCCs. Fig. 2 depicts a digraph constructed from (21). The exciter variables R_{fi} , V_{Ri} , and E_{fdi} form an SCC, but not a root SCC, since there is an incoming edge. Conversely, E'_{qi} , E'_{di} , ω_i and δ_i form a root SCC. A measurement depending on *at least* one variable in the root SCC makes the system observable, cf. (9). Clearly, I_{Di} and I_{Qi} depend on three variables within the root SCC, satisfying the condition for observability.

We verify this result using the $\mathcal{L}$ approach. We construct an observability matrix by computing successive Lie derivatives of the output $\mathbf{y} = [I_{Di} \ I_{Qi}]^T$. The $\mathcal{L}$ approach yields a full-rank observability matrix of dimension 7, confirming that the system is observable in the sense of Definition 3. Table II compares the average computation time required to check observability using the $\mathcal{L}$ and graph approaches.

Case 2: Centralized DSE

We perform a Kron reduction [14] of (10)–(13) to eliminate its algebraic constraints. This yields $\mathbf{x} = [\mathbf{x}_1^\top \mathbf{x}_2^\top \mathbf{x}_3^\top]^\top$, $\mathbf{u} = [\mathbf{u}_1^\top \mathbf{u}_2^\top \mathbf{u}_3^\top]^\top$, and $\mathbf{y} = [\mathbf{y}_1^\top \mathbf{y}_2^\top \mathbf{y}_3^\top]^\top$, where

$$\mathbf{x}_i = [E'_{qi} \ E'_{di} \ \delta_i \ \omega_i \ E'_{fdi} \ R_{fi} \ V_{Ri}]^\top, \\ \mathbf{u}_i = [T_{Mi} \ V_{refi}]^\top, \text{ and } \mathbf{y}_i = [V_{Di} \ V_{Qi}]^\top.$$

The updated model is written in the simplified form:

$$\begin{aligned} \dot{E}'_{qi} &= f'_{1i}(\{E'_{qj}, E'_{dj}, \delta_j, E'_{fdi}\}_{j=1}^3, \{\}) \\ \dot{E}'_{di} &= f'_{2i}(\{E'_{qj}, E'_{dj}, \delta_j\}_{j=1}^3, \{\}) \\ \dot{\delta}_i &= f'_{3i}(\{\omega_i\}, \{\}) \\ \dot{\omega}_i &= f'_{4i}(\{E'_{qj}, E'_{dj}, \delta_j, \omega_i\}_{j=1}^3, \{T_{Mi}\}) \\ \dot{E}'_{fdi} &= f'_{5i}(\{E'_{fdi}, V_{Ri}\}, \{\}) \\ \dot{R}_{fi} &= f'_{6i}(\{E'_{fdi}, R_{fi}\}, \{\}) \\ \dot{V}_{Ri} &= f'_{7i}(\{E'_{fdi}, R_{fi}, V_{Ri}\}, \{V_{refi}\}), \quad i = \{1, 2, 3\}. \end{aligned} \quad (22)$$

A digraph is less amenable to visualization in this case due to the strong coupling among state variables. Therefore, we only show the adjacency matrix, see Table III. Algorithm 1 identifies three SCCs, $\{V_{R1}, R_{f1}, E'_{fd1}\}$, $\{V_{R2}, R_{f2}, E'_{fd2}\}$, and $\{V_{R3}, R_{f3}, E'_{fd3}\}$; and one root SCC, $\{E'_{q1}, E'_{d1}, \delta_1, \omega_1, E'_{q2}, E'_{d2}, \delta_2, \omega_2, E'_{q3}, E'_{d3}, \delta_3, \omega_3\}$. Measuring at least one variable in each root SCC yields observable system. For instance, measuring

$$V_{Di} = g_3(E'_{qi}, E'_{di}, \delta_i) \text{ and } V_{Qi} = g_4(E'_{qi}, E'_{di}, \delta_i),$$

$i = \{1, 2, 3\}$, yields observability. The $\mathcal{L}$ approach returns a full rank(O) = 21 (3 generators, 7 state variables per generator) for the same measurement setup and, on average, requires 2 hours to compute. The average computation time of the graph approach is less than 5 seconds, an average $1440 \times$ speedup even in such a small test system. See Table II.

To further demonstrate the scalability of the graph method, we analyze the IEEE 39-bus, 10-machine system using a centralized DSE formulation, which involves 70 state variables. Thanks to the graph approach's linear-time complexity, 10 non-root SCCs and 1 root SCC are identified in roughly 30 seconds on average. In this case, the $\mathcal{L}$ approach fails to compute within polynomial time.

V. CONCLUSIONS AND FUTURE WORK

Current methods for observability analysis do not scale, especially in *centralized* dynamic state estimation. Even in the small test system we examined, with only three generators and twenty-one state variables, our graph-based approach yields a substantial reduction in computation time. The proposed method will facilitate new advances in power system dynamic state estimation, including new optimization formulations for synchrophasor measurement placement. Future research will build on the proposed graph methodology to develop a metric for quantifying observability strength, thereby enabling evaluation of sensitivity to parameter changes and measurement noise. This metric will help address exceptional cases where structural observability exists but local observability is lost due to parameter values or operating points. Additionally, the study will investigate structure-preserving models of power system networks that incorporate inverter-based resources.

TABLE II
AVERAGE EXECUTION TIMES FOR CASES 1 AND 2.

Case	Approach	$\mathcal{L}$	Graph
1: Decentralized DSE		5 minutes	1 second
2: Centralized DSE, 9-Bus system		2 hours	5 seconds
2: Centralized DSE, 39-Bus system		—	30 seconds

TABLE III
ADJACENCY MATRIX FOR THE CENTRALIZED DSE CASE

	E'_{d1}	E'_{d2}	E'_{d3}	E'_{q1}	E'_{q2}	E'_{q3}	δ_1	δ_2	δ_3	ω_1	ω_2	ω_3	E'_{fd1}	E'_{fd2}	E'_{fd3}	R_{f1}	R_{f2}	R_{f3}	V_{R1}	V_{R2}	V_{R3}
E'_{q1}	1	1	1	1	1	1	1	1	1												
E'_{d1}	1	1	1	1	1	1	1	1	1												
E'_{q2}	1	1	1	1	1	1	1	1	1												
E'_{d2}	1	1	1	1	1	1	1	1	1												
E'_{q3}	1	1	1	1	1	1	1	1	1												
E'_{d3}	1	1	1	1	1	1	1	1	1												
δ_1																					
δ_2																					
δ_3																					
ω_1	1	1	1	1	1	1	1	1	1	1	1	1									
ω_2	1	1	1	1	1	1	1	1	1	1	1	1									
ω_3	1	1	1	1	1	1	1	1	1	1	1	1									
E'_{fd1}													1								
E'_{fd2}														1							
E'_{fd3}															1						
R_{f1}																1					
R_{f2}																	1				
R_{f3}																		1			
V_{R1}																			1		
V_{R2}																				1	
V_{R3}																					1

REFERENCES

- [1] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Application*. CRC Press, 2004.
- [2] J. Zhao *et al.*, "Power system dynamic state estimation: Motivations, definitions, methodologies, and future work," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3188–3198, 2019.
- [3] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Trans. Autom. Control*, vol. 22, no. 5, pp. 728–740, 1977.
- [4] B. de Jager, "The use of symbolic computation in nonlinear control: is it viable?" *IEEE Trans. Autom. Control*, vol. 40, no. 1, pp. 84–89, 1995.
- [5] A. Rouhani and A. Abur, "Observability analysis for dynamic state estimation of synchronous machines," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3168–3175, 2017.
- [6] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Prentice Hall, 1997.
- [7] A. J. Krener and K. Ide, "Measures of unobservability," in *IEEE Conf. Decis. Control*, 2009, pp. 6401–6406.
- [8] J. Qi, K. Sun, and W. Kang, "Optimal PMU placement for power system dynamic state estimation by using empirical observability Gramian," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2041–2054, 2015.
- [9] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Observability of complex systems," *Proc. Natl. Acad. Sci.*, vol. 110, no. 7, pp. 2460–2465, 2013.
- [10] M. Sharir, "A strong-connectivity algorithm and its applications in data flow analysis," *Comput. Math. Appl.*, vol. 7, no. 1, pp. 67–72, 1981.
- [11] M. Netto, V. Krishnan, Y. Zhang, and L. Mili, "Measurement placement in electric power transmission and distribution grids: Review of concepts, methods, and research needs," *IET Gener. Transm. Distrib.*, vol. 16, no. 5, pp. 805–838, 2022.
- [12] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Supporting information for "observability of complex systems"," Supplementary Material to *Proc. Natl. Acad. Sci.* 110 (7): 2460–2465, 2013.
- [13] A. K. Singh and B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 794–804, 2014.
- [14] F. Dörfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Trans. Circuits Syst. I: Regul. Pap.*, vol. 60, no. 1, pp. 150–163, 2013.